IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

THOMSON REUTERS ENTERPRISE CENTRE GMBH and WEST PUBLISHING CORPORATION,	))))	
	)	C.A. No. 20-613-SB
Plaintiffs/Counterdefendants,	)	
	)	**JURY TRIAL DEMANDED**
v.	)	
	)	
ROSS INTELLIGENCE INC.,	)	**PUBLIC VERSION**
	)	
Defendant/Counterclaimant.	)	

ROSS INTELLIGENCE INC.'S MEMORANDUM OF POINTS AND AUTHORITIES IN OPPOSITION TO THOMSON REUTERS' RENEWED MOTION FOR PARTIAL SUMMARY JUDGMENT ON FAIR USE

OF COUNSEL:

Warrington S. Parker III
Joachim B. Steinberg
Jacob Canter
CROWELL & MORING LLP
3 Embarcadero Ctr., 26th Floor
San Francisco, CA 94111
Tel: (415) 986-2800

Keith J. Harrison
Crinesha B. Berry
CROWELL & MORING LLP
1001 Pennsylvania Avenue NW
Washington, DC 20004
Tel: (202) 624-2500

David E. Moore (#3983)
Bindu A. Palapura (#5370)
Andrew L. Brown (#6766)
POTTER ANDERSON & CORROON LLP
Hercules Plaza, 6th Floor
1313 N. Market Street
Wilmington, DE 19801
Tel: (302) 984-6000
dmoore@potteranderson.com
bpalapura@potteranderson.com
abrown@potteranderson.com

Attorneys for Defendant/Counterclaimant
ROSS Intelligence, Inc.

Dated: October 30, 2024
11852645 / 20516.00001
Public Version Dated: November 6, 2024

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Page(s)

Cases

Statutes

Other Authorities

INTRODUCTION AND SUMMARY OF ARGUMENT

ROSS set out to reduce the cost of legal research and increase the accessibility of the law by adapting artificial intelligence ("AI") technology to legal search engines. Contrary to the legally and factually incorrect assertions in Plaintiffs' motion for summary judgment, the method ROSS used was transformative and constitutes fair use. Plaintiffs' motion should be denied.

To the extent there was any use of copyrighted material, that use was transformative because ROSS translated the words of questions and judicial opinions into mathematical relationships to create a tool that, based on a legal question, output relevant judicial passages. The way that words relate to one another mathematically is not the expression, meaning, or message of Westlaw Content. Any expression, meaning, or message of Westlaw Content is removed through this process. And the output of ROSS's tool did not contain or depend on any of the claimed copyrighted materials at issue in this case.

These material facts regarding ROSS's transformative process and transformative output are undisputed. Because these facts are undisputed, Plaintiffs' summary judgment motion should be denied and ROSS's summary judgment motion on fair use should be granted.

It is undisputed that neither the mathematical process employed to create the search engine output nor the search engine output itself contains or uses any copyrighted material. It is undisputed that there is no copyrighted material in the source code. This is transformative use.

Additionally, to the extent ROSS accessed any of Plaintiffs' copyrights, it only did so to access non-copyrightable information. That non-copyrightable information includes the judicial passages on Westlaw and the non-expressive characteristics of the words in the text.

As for the other factors, the copyrights at issue to the extent creative are thin copyrights. To the extent implicated, the scope of use was minuscule compared to the entirety of the claimed scope of the copyrighted materials.

Rather than dispute these facts, Plaintiffs attempt to avoid a conclusion of fair use by factual generalization and the urging of an incorrect legal standard.

Plaintiffs contend that this Court should begin its analysis with *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023). However, this case is not apposite. Rather, *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1 (2021), and the Ninth Circuit's decisions in *Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992) and *Sony Computer Ent., Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000) provide the correct analytical framework. In those cases, the Supreme Court and Ninth Circuit correctly held that the kind of intermediate, transformative copying that ROSS did is fair use. As another example, Plaintiffs also claim that "sweat of the brow" is an appropriate legal consideration despite its rejection by *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991).

As for the facts, three examples show how Plaintiffs want to deflect from the undisputed material facts. Example one: Plaintiffs do not dispute the actual tiny amount used of the claimed copyrights at issue. They seek to avoid this by claiming that the "heart" of the copyrighted materials were used. D.I. 673 (herein, "Brief") at 36-40. Plaintiffs do not explain how that could be or why the particular pieces of Westlaw Content taken are the heart of their large work. *Id*.

Example two: Plaintiffs claim that the market implicated for factor four is the AI training data market even though ROSS did not participate in this market and had no intent to do so. *Id*. at 16-20. There is also no evidence that Plaintiffs' ability to enter that market was impaired, that Plaintiffs ever wanted to enter that market, or that Plaintiffs ever will.

Example three: Plaintiffs claim that ROSS sought LegalEase to avoid the ban they knew prohibited ROSS's use of Westlaw. In Plaintiffs' words, ROSS "contracted with LegalEase to provide training content . . . that copied Westlaw Content." Brief at 10. And this according to

Plaintiffs is bad faith precluding fair use. First, the fact is wrong. ROSS did not direct LegalEase to use Westlaw and LegalEase did not use Westlaw exclusively. Second, the argument is legally irrelevant. Bad faith is not an issue. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 585 n.18 (1994); *Google LLC*, 593 U.S. at 32-33. Last, as Plaintiffs portray the facts, ROSS did nothing more than what Google did in *Google LLC*. ROSS sought a license. When that did not happen, it obtained the material a different way.

When analyzed under the correct line of cases, the undisputed material facts show that all four fair use factors favor ROSS. This Court should deny Plaintiffs' motion for summary and judgment and grant ROSS's motion for summary judgment on fair use.

STATEMENT OF FACTS

Plaintiffs' Motion for Summary Judgment on Fair Use misconstrues material facts and focuses on immaterial facts. Critically, it also ignores material, including undisputed material facts related to the transformative steps ROSS took to create its AI tool.

I.   ROSS' FOUNDING AND PURPOSE

Jimoh Ovbiagele understood at a personal-level how expensive and ineffective legal research products could make it difficult for real people obtain legal services. Supplemental Declaration of Jacob Canter ("Suppl. Canter Decl.") Ex. 1, J. Ovbiagele Dep. Tr. at 38:17-25; Suppl. Ovbiagele Decl. Ex. 2, ROSS-010129852 at -852. So, in 2014 he teamed up with Andrew Arruda to start ROSS. Suppl. Canter Decl. Ex. 2, A. Arruda Dep. Tr. at 11:21-12:4. ROSS's goal was to build something new—a tool that "allow[s] lawyers to do a lot more with less" and brings us closer to a day "when everyone can afford a lawyer." Supplemental Declaration of Jimoh Ovbiagele ("Suppl. Ovbiagele Decl.") Ex. 3, ROSS-001397429 at -431.

ROSS's plan was to use state-of-the-art AI technology to achieve this goal. Suppl. Canter Decl. Ex. 1, J. Ovbiagele Dep. Tr. at 41:5-42:1. Unlike existing tools, ROSS's technology would

3

be "free of human intermediated content." Ovbiagele Declaration, D.I. 680 ("Ovbiagele Decl."),

¶ 5; Suppl. Ovbiagele Decl. Ex. 4, ROSS-010271831 at -831 (ROSS "aims to recognize and

extract answers directly from the law using machine learning. It has always been our view that

the editorialized content approach . . . is the old legal research playbook. We chose a new way.").

Using ROSS's AI, a user could ask a legal question in plain English and receive, in

response, direct answers from the judicial opinions. Suppl. Ovbiagele Decl. Ex. 5, ROSS-

003487253 at -290; *id*. Ex. 6, ROSS-023179081 at -087; *id*. Ex. 7, ROSS-003515560 at -560; *id*.

Ex. 8, ROSS-009542437 at -437; *id*. Ex. 3, ROSS-001397429 at -429. By making legal research

less time-consuming, the price for legal services would drop. *Id*. Ex. 3, ROSS-01397429 at -431;

id. Ex. 2, ROSS-010129852 at -852.

II.   WHY ROSS RETAINED LEGALEASE AND LEGALEASE'S ACTIVITIES

ROSS needed a large and diverse body of legal questions and answers that were verbatim

quotes from judicial opinions to train its AI. Ovbiagele Decl. ¶ 30; Suppl. Canter Decl. Ex. 3,

TR-0000568 at -569. ROSS obtained these question-answer pairs in memos. Suppl. Ovbiagele

Decl. Ex. 9, ROSS-000000001. ROSS hired LegalEase to do this work ("bulk memo project").

Suppl. Canter Decl. Ex. 1, J. Ovbiagele Dep. Tr. at 116:12-16; *id*. Ex. 3, TR-0000568. LegalEase

hired a subcontractor, Morae (collectively, "LegalEase"), to assist with the memos. *Id*. Ex. 4, T.

Hafeez Dep. Tr. at 88:13-18; *id*. Ex. 5, C. Cahn Dep. Tr. at 25:2-28:2; i*d*. Ex. 6, TR-0039943.

Plaintiffs contend that ROSS licensed with LegalEase to access Westlaw Content without

Plaintiffs' knowledge. Brief at 9-10. However, ROSS did *not* direct the use of Westlaw by

LegalEase. Suppl. Canter Decl. Ex. 7, T. Whitehead Dep. Tr. at 43:2-5, 189:2-6. And in fact,

LegalEase did *not* use Westlaw exclusively. *Id*. Ex. 4, T. Hafeez Dep. Tr. at 61:12-25, 224:15-

225:6; *id*. Ex. 7, T. Whitehead Dep. Tr. at 60:10-19, 133:5-21, 189:13-22; *id*. Ex. 8,

LEGALEASE-00078065 at -067-76. When LegalEase did use Westlaw, it paid for Westlaw

pursuant to a license. *Id*. Ex. 4, T. Hafeez Dep. Tr. at 321:3-19.

Plaintiffs also contend that ROSS contracted with LegalEase so that LegalEase could create training data that "copied Westlaw Content." Brief at 10. That cannot be so given the facts immediately above. Nor do Plaintiffs cite evidence to support this claim.

Plaintiffs cite to their Exhibits 36 and 44 for this claim. Exhibit 36 is a "Best Practices Guide" that states that headnotes "are proprietary and you should not copy" them. Suppl. Canter Decl. Ex. 8, LEGALEASE-0078065 at -068. The exhibit also references key numbers as a means to locate the text of judicial opinions. *Id*. It says nothing about copying key numbers and makes no reference to any other type of "Westlaw Content." *Id*. Exhibit 44 is an email that makes no reference to Westlaw Content at all. Suppl. Ovbiagele Decl. Ex. 10, ROSS-000304769.

Finally, when ROSS did receive materials that contained Westlaw Content—500 judicial opinions—from LegalEase, ROSS responded by asking LegalEase to ▮▮▮▮▮▮▮▮▮▮▮ ▮▮▮▮▮▮▮▮ such as ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮" because ROSS *did not* want it. Suppl. Ovbiagele Decl. Ex. 11, ROSS-000197949 at -949. Those materials were ▮▮▮▮ and ROSS only wanted ▮▮▮▮▮▮▮▮ *Id*. ROSS said that this ▮▮▮▮ material was ▮▮▮▮ and ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮ ▮▮▮▮▮▮▮▮ Suppl. Canter Decl. Ex. 9, T. van der Heijden Dep. Tr. at 135:23-136:23.

III. HOW LEGALEASE PREPARED THE MEMOS

To prepare the memos, LegalEase used Westlaw and LexisNexis. Suppl. Canter Decl. Ex. 4, T. Hafeez Dep. Tr. at 61:12-25, 224:15-225:6; *id*. Ex. 7, T. Whitehead Dep. Tr. at 60:10-19, 133:5-21. The questions in the memos were based on text from judicial opinions. *Id*. Ex. 8, LEGALEASE-00078065 at -067-76.

If one at LegalEase did copy headnotes verbatim, he (1) violated the Best Practices Guide referenced by Plaintiffs, *see* Exhibit 36 at LEGALEASE-0078065 at -068; and (2) failed the

purpose of providing the memos ████████████████████████████████████

████████████████████████ *Id*. Suppl. Canter Decl. Exs. 8, 5, C. Cahn Dep. Tr. at 64:22-65:13.

As to the first point, LegalEase's guidelines directed them to not copy headnotes, stating

that "the head notes [sic] are proprietary and you should not copy paste them in the question." *Id*.

Ex. 8, LEGALEASE-00078065 at -068. LegalEase's leadership instructed its contract attorneys

to not copy headnotes. *Id*. Ex. 4, T. Hafeez Dep. Tr. at 63:6-25, 78:1-17, 84:20-85:9, 225:20-

228:25; *id*. Ex. 10, R-LEGALEASE-00050673 at -673; *id*. Ex. 5, C. Cahn Dep. Tr. 62:21-63:21,

64:22-65:13. LegalEase only sought to ██████████████████ and stated that ████████

████████████████████████████████ *id*. Ex. 11, TR-0047926 at -926.

Each memo had four to six verbatim quotes from a judicial opinion. Suppl. Ovbiagele

Decl. Ex. 9, ROSS-000000001; Suppl. Canter Decl. Ex. 5, C. Cahn Dep. Tr. at 54:15-55:21.

Each quote was given a relevancy rating—great, good, topical or irrelevant. Suppl. Ovbiagele

Decl. Ex. 9, ROSS-000000001; TR-0000568 at -568-69. A "great" quote "contain[ed] an answer

to all essential elements" of the legal question. Suppl. Canter Decl. Ex. 3, TR-0000568 at -568

(section 5.3). A "good" quote did not provide a complete answer to the question. *Id*. Ex. 3, TR-

0000568 at -568 (section 5.3). A topical quote did not answer any of the question's essential

elements. *See id*. Ex. 8, LEGALEASE-000078065 at -078. An irrelevant quote "completely

missed the mark." *See id*. Ex. 8, LEGALEASE-000078065 at -078.

Just looking at the face of the memos shows that they do not contain any Westlaw

Content. Suppl. Ovbiagele Decl. Ex. 9, ROSS-000000001. Nothing in the memos represents

West created legal analysis, summaries or syntheses of cases, or the key number system. *Id*.

Plaintiffs contend that every memo question was copied from a West headnote. Brief at

10. Plaintiffs cite Exhibits 36, 37, and 39 for this proposition. *Id*. These documents are "guides"

and "protocols" that LegalEase and Morae used as models for preparing bulk memos. *See* Suppl. Canter Decl. Ex. 8, LEGALEASE-000078065 (the "Best Practices Guide"); *id*. Ex. 12, LEGALEASE-00093066 (the "Project Rose Project Protocol"); *id*. Ex. 13, R-LEGALEASE-00189134 (images depicting how to prepare a memo).

These guides and protocols do not show that every memo question was copied from a West headnote. *See id*. Exs. 8, 12, 13. The Best Practices Guide in fact states the opposite: that employees "should not copy paste" headnotes when drafting memo questions because headnotes are "proprietary." *See id*. Ex. 8, LEGALEASE-000078065 at -068. It also includes a description for how to draft a memo using *Lexis*. *See id*. Ex. 8, LEGALEASE-000078065 at -072-76.

Plaintiffs also contend that every judicial passage in a memo was selected by an attorney-editor to correspond to a headnote. Brief at 10. This is irrelevant and unsupported. Plaintiffs cite Exhibits 4 and 86 for this claim. *Id*. Exhibit 4 is testimony from Morae that describes how Morae drafted memo *questions*. Suppl. Canter Decl. Ex. 5, C. Cahn. Dep. Tr. at 195:14-19. Exhibit 86 is an email between LegalEase employees that states that starting with a headnote is one way to locate great passages. *See* Suppl. Canter Decl. Ex. 14, TR-0073545 at -545. This was one way, but not the only way. *Id*. Ex. 5, C. Cahn Dep. Tr. at 68:24-69:15 (another way to locate passages was by ███████████████████ 180:18-181:16 ("███████████████████ ████████████████████████████ *id*. Ex. 4, T. Hafeez Dep. Tr. at 81:3-82:2.

Also, Exhibit 86 says nothing about how LegalEase located judicial passages to receive a good, topical, or irrelevant label. *See id*. Ex. 14, TR-0073545. The Project Rose Protocol shows that LegalEase did not locate topical and irrelevant quotes based on headnotes. It explains that "[t]he easiest way to find topical cases is to look for cases within the cases cited[.]" *Id*. Ex. 12, LEGALEASE-00093066 at -070. And it states that "the easiest way to locate a good irrelevant

7

quote is to pick a key word from the question, search only that word in WestLaw/Lexis, [and] find a case with that word in the facts with little to no other words from the question[.]" *Id*.

Finally, as reflected in ROSS's opposition to Plaintiffs' Infringement Summary Judgment Motion, the claim that every question was copied from a headnote is demonstrably false. Some of the cases on which Plaintiffs rely to show "infringement" do not contain headnotes or the headnote does not correspond to the case identified by Plaintiffs. In many instances, there is no correlation between the question and the headnote at issue.[1]

IV.   MATERIAL AND UNDISPUTED FACTS AS TO ROSS'S TRANSFORMATIVE PROCESS—HOW ROSS USED THE QUESTION-ANSWER PAIRS TO TRAIN ITS AI ON NON-EXPRESSIVE CHARACTERISTICS OF THE TEXT

Plaintiffs ignore every fact about the development of ROSS's AI tool. These are material facts because they show that ROSS removed any expression, meaning, or message from the Westlaw Content through the training process and they show that ROSS's tool does not use any expression, meaning, or message from the Westlaw Content.

A.   Preparing the Question-Answer Pairs for Computational Analysis

ROSS first ███ for processing. Ovbiagele Decl. ¶¶ 16-17; Declaration of Joseph Marks, D.I. 689 ("Marks Decl.") ¶ 12. As part of this process, ROSS had quality control mechanisms in place: ████████████

███

██████████████████████████████. Ovbiagele Decl. ¶¶ 7-10, 14, 16-17;

Marks Decl. ¶ 12. Due to these quality control steps ROSS eliminated any arrangement or organization from the key number system built into the memos or pairs. Marks Decl. ¶ 12.

---

[1] ROSS's opposition to Plaintiffs motion for summary judgment on copyright infringement is being filed at the same time as this brief.

Next, ROSS randomly allocated ███ of the question-answer pairs for use as training data and ███ of the question-answer pairs for testing and validation. Ovbiagele Decl. ¶¶ 6, 32; Marks Decl. ¶ 12. At this step as well, any arrangement or organization from the key number system that was built into the memos or question-answer pairs was eliminated. Marks Decl. ¶ 12.

Then, the remaining pairs allocated for testing were prepared for a computational analysis of their non-expressive characteristics through processes called *tokenization* and *lemmatization*. Ovbiagele Decl. ¶ 18; Marks Decl. ¶¶ 18-19. Tokenization removes "stop" words (like "is," "to," or "be") that are useful for human comprehension but irrelevant for computational analysis. *Id*. And lemmatization changes the grammatical inflection of all the words to be the same with respect to number, tense, and voice (so "swam" and "swum" both become "swim") which also makes the text harder for a human to understand but easier for a computer to analyze. *Id*.

B.   The Featurization of the Question-Answer Pairs

The question-answer pairs that were tokenized and lemmatized next were transformed via an algorithm in ROSS's source code into numerical representations of abstract characteristics and features about the text. Ovbiagele Decl. ¶¶ 4, 19, 22-30, 28-29; Marks Decl. ¶¶ 13-16, 20-24, 29-36. Each question-answer pair that reached this step was transformed into ███████ different numbers, all representing different information about the syntax or structure of the pairs. Ovbiagele Decl. ¶¶ 19, 22-30, 28-29; Marks Decl. ¶¶ 13-16, 20-24, 29-36. This step— generating numbers that represent non-expressive features about the text—is called *featurization*. Ovbiagele Decl. ¶ 20; Marks Decl. ¶ 20.

ROSS chose which non-expressive features about the text to generate at this step. Ovbiagele Decl. ¶ 19; Marks Decl. ¶ 20. One example of a feature is called ██████████ ██ Ovbiagele Decl. ¶ 19.20; Marks Decl. ¶ 20.20. Another feature is called ██████████████████████

9

███

███

Ovbiagele Decl. ¶ 19.14; Marks Decl. ¶ 20.14. Another feature is called ██████████

███

███

█████████████████. Ovbiagele Decl. ¶ 19.24 to .27; Marks Decl. ¶ 20.24 to .27. Another

feature is called ██

███

████████ Ovbiagele Decl. ¶ 19.1; Marks Decl. ¶ 20.1.

The remaining features are the same as these. They also produce a number that represents

something non-expressive about the text—something unrelated to the Westlaw Content's

expression, meaning, or message. Ovbiagele Decl. ¶¶ 19, 22-30; Marks Decl. ¶¶ 20-24, 29-36.

Once the features were derived, ROSS did not use the text of the question-answer pairs

and that text was discarded. Ovbiagele Decl. ¶ 22; Marks Decl. ¶ 22. This is no different than

Plaintiffs' use of training data. Plaintiffs' witness Dr. Isabelle Moulinier stated that Plaintiffs' AI

tools also did not use the text of question-answer pairs after the training occurred. *See* Suppl.

Canter Decl. Ex. 15, I. Moulinier Dep. Tr. at 61:22-62:6 (once the AI tool was ████████████ it

was █████████████████████ and the text was no longer used or relevant).

C.   **Using the Numbers Representing Non-Expressive Features of the Text of the Question-Answer Pairs to Train ROSS's Artificial Intelligence**

The numerical features that ROSS generated were used by the "learning to rank"

component of ROSS's machine-learning model, called █████████████ to compute parameters

(or guideposts) that train the technology to predict whether a candidate judicial passage is

relevant to a natural language query. Ovbiagele Decl. ¶ 26; Marks Decl. ¶ 15. Through this

10

process, the features cause the learning-to-rank component of the model to change. Ovbiagele Decl. ¶ 26; Marks Decl. ¶¶ 15-16. The order, structure, or arrangement of the features and labels is irrelevant in this training process. Ovbiagele Decl. ¶¶ 22, 32-35; Marks Decl. ¶¶ 15-16, 20-22. ROSS's source code contained no Westlaw Content. Ovbiagele Decl. ¶ 33; Marks Decl. ¶ 13, 34.

D.   Ross Did Not Want or Use Plaintiffs' Legal Analysis

In addition to ignoring every fact about how ROSS developed its AI tool, Plaintiffs contend that "ROSS wanted *legal analysis*" to "teach its AI model how to answer legal questions" because it "served as an excellent synthesis and organization of the law[.]" Brief at 1, 9 (emphasis in original). Plaintiffs also contend that ROSS "used the Westlaw Content as summations and syntheses of factual and legal issues, selected and arranged in coherent categories, chosen and honed by Plaintiffs, within Plaintiffs' original West Key Number System" to train its algorithm. Brief at 28.

Even had ROSS received more than questions answered specifically by the words of judicial opinions such that it is appropriate to claim that ROSS received what Plaintiffs claim is analysis, synthesis, organization, selection, or summation from the Westlaw Content, all that was changed to mathematical equations reflecting the relationship of words. Statement of Facts § IV.A-C. There is no longer "legal analysis," summations, arrangements, categories, or syntheses of factual and legal issues—even were these valid descriptions of the copyrighted materials. *Id*.

Also, Plaintiffs misconstrue the facts it relies on to establish these propositions. Plaintiffs cite Exhibit 2 to attempt to show that ROSS wanted legal analysis. Brief at 9. But this quote from Mr. Arruda is unrelated. It only establishes that ROSS ███████████████████ ████████████████████████████████████ because the way to train an AI tool is by using question-answer pairs. Suppl. Canter Decl. Ex. 2, A. Arruda Dep. Tr. at 275:23-276:12.

Moreover, when Plaintiffs make this argument at page 28 of their brief, they cite to pages

11

8 through 10 and 16 as the basis for the claim that ROSS used the Westlaw Content "as

summations and syntheses of factual and legal issues" to train its algorithm, and that ROSS used

the key number system to train its algorithm. Brief at 28.

However, most of page 8 describes why Plaintiffs believe Westlaw Content is valuable,

which is irrelevant. *Id*. at 8. Pages 8-10 cite to no documents that support the notion that ROSS

used summations or syntheses of cases, or used the key number system, to train its algorithm.

Brief at 8-10. Instead, the documents referenced describe ROSS's competitive efforts and

business and marketing strategies[2]; that question-answer pairs are used to train the AI algorithm

and that ROSS sought caselaw that did not contain Westlaw Content for its pairs[3]; that ROSS

---

[2] Every citation in this paragraph to the Arruda transcript (Ex. 2) and the van der Heijden transcript (Ex. 9) relates to marketing, competitive efforts, or ROSS's founding. Suppl. Canter Decl. Ex. 2, A. Arruda Dep. Tr. at 11:18-12:4 (ROSS started in 2014); 39:1-3 (ROSS and Plaintiffs compete); 114:16-115:11 (ROSS and Plaintiffs compete and ROSS sought to grow in the market); *Id*. Ex. 9, T. van der Heijden Dep. Tr. at 55:21-56:20 (ROSS's intention was to build a tool that replaces Westlaw). The ROSS blog post (Pls.' Ex. 94) describes how natural language processing operates, makes no reference to Westlaw Content, and does not state that the technology functions by summarizing or synthesizing legal analysis or by relying on a method of organizing caselaw. *Id*. Ex. 16, ROSS Blog Post. The remaining evidence in this paragraph is three documents (Pls. Exs. 64, 70, 57), and each of these describe ROSS's marketing and business efforts. Suppl. Ovbiagele Decl. Ex. 12, ROSS-009501052 at -052 (bullet point titled "Business Motivation"); *id*. Ex. 19, ROSS-010099622 at -623 (a draft marketing blog post); *id*. Ex. 13, ROSS-003395895 (a pricing document).

[3] The citation to Jimoh Ovbiagele's transcript (Pls.'Ex. 13) establishes four propositions that state nothing about ROSS's tool synthesizing or summarizing legal or factual issues or about ROSS's tool using the key number system: the transcript states (i) that ROSS's product ██████ ████████████████████████████████████ ; (ii) that ROSS's machine learning model was trained to assign a relevance score to a question-answer pair; (iii) that ROSS's tool operated through a search engine that provided the machine learning model candidate judicial quotes and the user's question; and (iv) that ROSS wanted quotes from case law for its training data. Suppl. Canter Decl. Ex. 1, J. Ovbiagele Dep. Tr. at 48:23-50:9, 204:3-7. Plaintiffs' Exhibit 73 is a blog post drafted by ROSS that states the *opposite* of the conclusion Plaintiffs seek from this document. Suppl. Ovbiagele Decl. Ex. 4, ROSS-010271831 at -831 (stating that ROSS "never asked for or used proprietary data from Westlaw. Not only could we not map this proprietary data to the case law in our database, our technology does not attempt to generate editorialized answers; instead, it aims to recognize and extract answers directly from the law using machine learning."). The citation to the van der Heijden transcript (Pls.' Ex. 16) only

12

attempted to access Westlaw several years prior to the bulk memo project for reasons unrelated

to that project;[4] and that LegalEase created the memos.[5]

Similarly, none of the documents cited on page 16 state that ROSS used summations or

syntheses of cases or the key number system to train its AI. The citations to Dr. Moulinier's and

Dr. Alan Cox's transcripts establish that *Plaintiffs* use Westlaw Content to train their AI. Suppl.

Canter Decl. Ex. 15, I. Moulinier Dep. Tr. at 72:4-11; *id*. Ex. 18, Cox Dep. Tr. at 123:21-124:23.

The citations to Ms. Laurie Oliver's declaration and Mr. James Malackowski's transcript

establish that *Plaintiffs* consider Westlaw Content valuable. Declaration of Laurie Oliver, D.I.

256 ("Oliver Decl.") ¶ 13; Suppl. Canter Decl. Ex. 19, J. Malackowski Dep. Tr. at 42:3-24. And

the citation to Mr. Ovbiagele's transcript establishes that ROSS did not want LegalEase to sell

the bulk memos to another. Suppl. Canter Decl. Ex. 1, J. Ovbiagele Dep. Tr. at 167:24-170:19.

Finally, the quote on page 1 of Plaintiffs' brief has no support. Brief at 1.

V.   MATERIAL AND UNDISPUTED FACTS AS TO ROSS'S TRANSFORMATIVE OUTPUT—THE OUTPUT OF ROSS'S AI TOOL AND THE DIFFERENT OUTPUTS OF PLAINTIFFS' TWO AI TOOLS

After the training, ROSS's tool provides valuable legal research support. ROSS's AI tool

---

establishes that ROSS already had caselaw from Casemaker. Suppl. Canter Decl. Ex. 9, T. van der Heijden Dep. Tr. at 239:17-19. Paragraph 149 of Dr. Krein's report does not cite a single document and is just his opinion that ROSS wanted Westlaw Content. *Id*. Ex. 17, Expert Report of Jonathan Krein (Aug. 1, 2022) ¶ 149. And Plaintiffs' Exhibit 55 is a one-page slack where, in 2016, ROSS proposed accessing judicial text through Westlaw. Suppl. Ovbiagele Decl. Ex. 14, ROSS-003391075. The citation to the Arruda transcript is addressed *supra*.

[4]Four of these exhibits establish only that ROSS sought access to Westlaw's terms of use and bankruptcy materials in 2015 through the law firm Dentons, several years before the bulk memo project. Suppl. Ovbiagele Decl. Ex. 15, ROSS-023032254; Suppl. Canter Decl. Ex. 9, T. van der Heijden Dep. Tr. 101:19-103:5; Suppl. Ovbiagele Decl. Ex. 16, ROSS-003390233; *id.* Ex. 17, ROSS-003390772; Suppl. Canter Decl. Ex. 3, TR-0000568 at -571. A fifth exhibit establishes only that ROSS wanted to access Westlaw through a Thomson Reuters representative in 2015. Suppl. Ovbiagele Decl. Ex. 18, ROSS-003390563. And the last two exhibits are unrelated. Suppl. Canter Decl. Ex. 20, Resp. to Interrog. 11; *id*. Ex. 21, TR-0908413.

[5] *Infra* Statement of Facts § III.

outputs a rank-ordered collection of ***paragraphs from judicial opinions***. *See* Suppl. Ovbiagele Decl. Ex. 6, ROSS-023179081 at -087 ("ROSS returns the exact passages from the law . . ."); Marks Decl. ¶ 35 ("The only possible output from ROSS's tool was judicial passages."); Suppl. Ovbiagele Decl. Ex. 7, ROSS-003515560 at -560 (ROSS "read[s] and highlight[s] the answers in case law for you, taking you[] directly to the most important passages."). No Westlaw Content is present in this output. Ovbiagele Decl. ¶ 14 (passages from Casemaker opinions contain no Westlaw Content). This happens every time a ROSS user enters a query. Suppl. Ovbiagele Decl. Ex. 7, ROSS-003515560 at -560. If a user wants to know how a judge has answered a legal question, then ROSS's tool is a great fit. These facts about ROSS's technology are undisputed.

Plaintiffs contend that there is no difference between the outputs for ROSS's AI tool and Plaintiffs' artificial intelligence. Brief at 8, 27. This is a false statement.

As an initial matter, Plaintiffs fail to identify their AI tools accurately. Plaintiffs have two AI tools. Plaintiffs older AI tool is called WestSearch. Suppl. Canter Decl. Ex. 26, TR-0433738 at -738. Plaintiffs newer AI tool is called WestSearch Plus. *Id*. Ex. 22, TRCC-00618557 at-559. Plaintiffs fail to distinguish between WestSearch and WestSearch Plus even though their witnesses, experts, and documents make this distinction. *Id*. Ex. 15, I. Moulinier Dep. Tr. at 12:19-13:10; *id*. Ex. 23, Al-Kofahi Dep. Tr. at 17:6-18:4; *id*. Ex. 17, Expert Report of Jonathan Krein (Aug. 1, 2022) ¶¶ 76-78; *id*. Ex. 26, TR-0433738; *id.* Ex. 24, TR-0477495 at -500; *id*. Ex. 25, TRCC-0060822 at -830 (███████ became available in 2010 as part of ████████), -831 (" ███████ became available in 2018 as part of ████████ This distinction is important because it shows that Plaintiffs were not using headnotes whatsoever to train their AI technology before ROSS entered the market, and that even after ROSS entered the market Plaintiffs use of headnotes was limited to just their text. *Infra* Statement of Facts § VI.

In addition to not accurately identifying their own AI tools, Plaintiffs ignore material facts about the output for these tools, how they provide different legal research information, and how they are useful for different legal research purposes.

WestSearch outputs a list of *documents*. *See id*. Ex. 26, TR-0433738 at-738 (WestSearch ████████████████████████ (emphasis added); *id*. at -741; *id*. ████████████████ ██ (emphasis added); *id*. Ex. 27, TR-0456869 at -869 ("the most relevant *documents* are placed on top of the list") (emphasis added); *id*. Ex. 15, I. Moulinier Dep. Tr. at 13:17-14:5 ████████████████ ██████████████████████████████████████ (emphasis added). These documents come from twelve categories of ██████████████ on Westlaw, which includes cases, statutes, treatises, law reviews and bar journals, and legal newspapers. *See id.* Ex. 26, TR-0433738 at -739. And these documents include Westlaw Content, which the ROSS search engine could not return and did not return. *See id.* Ex. 26, TR-0433738 at -740. If a user wants a breadth of potential documents that could all potentially answer a legal question, and if the user has time to search through these various documents for the best fit, then WestSearch is a fine choice. *See id.* Ex. 26, TR-0433738 at -738 (WestSearch ████████████████████ and ██ *id.* Ex. 27, TR-0456869 at -869 ████████████████████████).

WestSearch Plus outputs *headnotes*. *id*. Ex. 28, TR-0908443 at -443 ("██████████ ██ *id*. Ex. 17, Expert Report of Jonathan Krein (Aug. 1, 2022) ¶ 86 ████████████████████████████ ████████████████; *id*. Ex. 15, I. Moulinier Dep. Tr. at 13:17-14:9, 14:25-15:5 ████████████ ████████████████████████████). It is not a list of headnotes, just ██████████ *id*. Ex. 15,

15

I. Moulinier Dep. Tr. at 13:17-14:9. Also, WestSearch Plus ███████ does not operate in

response to a user query, and this '███████[s]" customers. *id*. Ex. 29, TRCC-00988301 at -313

██

██

██

████████ (emphasis in original). If a user wants to know what a headnote states in response to

a question—even if it is not the user's specific question—then WestSearch Plus is the tool to use.

	What is the Output?	Does Output Contain Westlaw Content?	How Often Does It Respond to Queries?
ROSS's AI Tool	Rank-Ordered Set of Individual Paragraphs from Judicial Opinions (Suppl. Canter Decl. Ex. 6 at -087)	No (Ovbiagele Decl. ¶ 14)	Always (Suppl. Canter Decl. Ex. 7 at -560)
WestSearch	Rank-Ordered Set of Documents: cases, statutes, treatises, articles, etc. (Suppl. Canter Decl. Ex. 26 at -739)	Yes (Suppl. Canter Decl. Ex. 26 at -740)	Always (Suppl. Canter Decl. Ex. 26 at -738)
WestSearch Plus	A ██████ of headnotes (Suppl. Canter Decl. Ex. 15 at 13:17-14:9)	Yes (Suppl. Canter Decl. Ex. 15 at 14:25-15:5)	██████ Does Not (Suppl. Canter Decl. Ex. 29 at -313)

VI. HOW WESTSEARCH, WESTSEARCH PLUS, AND ROSS'S AI TOOL WERE TRAINED AND DIFFERENCES BETWEEN THE TRAINING OF THE THREE

Plaintiffs contend that WestSearch and WestSearch Plus were trained on Westlaw

Content in 2010, "before" ROSS trained its AI. Brief at 8, 16 (emphasis removed).

This cannot be the case for WestSearch Plus, because WestSearch Plus was introduced

into the market in 2018. Suppl. Canter Decl. Ex. 22, TRCC-00618557 at-559. This is eight years

after 2010 and four years after ROSS entered the market. *Id*. Ex. 2, A. Arruda Dep. Tr. at 11:18-

12:4. Also, WestSearch Plus was *only* trained on the text of headnotes as the answer and not as

the questions in training data. *Id*. Ex. 15, I. Moulinier Dep. Tr. at 77:23-78:4 (Headnotes are ████

██

████████); *Id*. Ex. 28, TR-0908443 at -443.

16

It is true that WestSearch was last trained in 2011. *Id*. Ex. 15, I. Moulinier Dep. Tr. at

71:22-72:25, 88:14-17. But the *only* Westlaw Content that Dr. Moulinier stated WestSearch was

trained on is ███████████ *Id*. Ex. 15, 126:11-128:6. Dr. Moulinier admitted that ██████████

██████████████████████████████████ WestSearch. *See id*. Ex. 15, 76:21-77:16.[6]

Entire documents were the answers for WestSearch training data—not Westlaw Content—which

is why WestSearch training data is ████████████████ *Id*. Ex. 15, 74:11-75:13.

And there is no evidence that WestSearch or WestSearch Plus were trained on the "links"

between cases and headnotes which Plaintiffs now focus so heavily upon in their argument. Brief

at 27, 31, 34. Even assuming that ROSS used some type of Westlaw Content for AI training, the

evidence confirms that ROSS used *different* Westlaw Content for training its AI than Plaintiffs

used to train WestSearch and WestSearch Plus.

Recognizing this problem, Plaintiffs focus on a higher level of generality: that Plaintiffs

used Westlaw Content "to provide relevant law in response to user queries" and that ROSS also

used Westlaw Content "to final relevant answers to legal research questions." Brief at 28. But

this higher level of generality ignores the important differences already described herein.

VII.   LICENSING OF WESTLAW CONTENT

Plaintiffs have unequivocally stated that at no time will they ever license Westlaw

Content as AI training data. Plaintiffs stated this in their responses to requests for admission and

interrogatories. Suppl. Canter Decl. Ex. 30, Resp. to Req. for Admis. 72 (Plaintiffs have a

█████████████████████████████████████ *id*. Ex. 31, Resp. to Interrog. 21

---

[6] Dr. Moulinier stated that headnotes were not used to train ███████████ because the AI tool
for Westlaw Next is WestSearch. *See* Suppl. Canter Decl. Ex. 26, TR-0433738; *id*. Ex. 23, Al-
Kofahi Dep. Tr. at 17:6-20 ████████████████████████
████████████████████

(Plaintiffs ███████████████████████████████ And Plaintiffs' declarant Ms.

Oliver stated that "Plaintiffs have a policy against selling their content to competitors for any

purpose, including training artificial intelligence algorithms." Oliver Decl. ¶ 13. It is not

reasonable to assume that Plaintiffs would license Westlaw Content (editorial enhancements to

legal materials) to companies that are not in the legal research marketplace. Declaration of Alan

Cox, D.I. 681 ("Cox Decl.") ¶¶ 11-12, 19-21. Moreover, ROSS never sold or licensed its memos

and never attempted to do so either. Ovbiagele Decl. ¶¶ 41-44; Cox Decl. ¶¶ 13, 25.

Nonetheless, Plaintiffs contend that ROSS appropriated Plaintiffs' chance to license

Westlaw Content as training data "at a strategically advantageous time and under suitable terms."

Brief at 17. Plaintiffs later argue that "[w]hether Plaintiffs actually participated in [the market to

license Westlaw Content as AI training data] is not legally relevant" because "they could

reasonably decide to participate" at a later time. Brief at 20. However, there still remains no

evidence that Plaintiffs ever did participate or ever would participate, or even that ROSS took

away Plaintiffs' ability to consider this participation at some future time.

ROSS, for its part, has never intended to sell or license and in fact has never sold or

licensed any Westlaw Content (or any content) as AI training data. Ovbiagele Decl. ¶¶ 41-44.

ROSS also has no intention to do so in the future. *Id*.

VIII.   PLAINTIFFS WESTLAW CONTENT

Plaintiffs state that Westlaw Content is valuable because it is comprehensive and

accurate. This is in their marketing. Suppl. Canter Decl. Ex. 32, TR-0179830 at -830 ("*Each*

legal issue in a case . . . is identified, summarized in a headnote, and assigned a topic and key

number"); *id*. Ex. 33, TR-0179843 at -843 (key number system "is comprehensive and precise").

This has also been stated by Plaintiffs' witness Mr. Andrew Martens, who could not state what

the value of a single headnote was because ███████████████████████████ and

18

instead ██████████████ where the ████████ of headnotes and key numbers is "█████ ████████████████████████." *Id*. Ex. 34, A. Martens Dep. Tr. at 224:15-226:1.

The purpose of Westlaw Content is to assist lawyers in finding judicial opinions for legal research. Declaration of Richard Leiter, D.I. 687 ("Leiter Decl.") ¶¶ 5-6 10, 13-22, 25; Cox Decl. ¶ 8. For this reason, Plaintiffs market that they hire attorneys to develop Westlaw Content that is accurate. *See* Suppl. Canter Decl. Ex. 33, TR-0179843 at -843; *id*. Ex. 35, TR-0432528 at -532.

Headnotes ██ ████████████." Suppl. Canter Decl. Ex. 36, L. Oliver Dep. Tr. at 32:21-25; *see also id.* Ex. 37, TR-0002864 at -872 (each point of law gets a headnote and ████████████████ ████████████████); *id.* Ex. 38, TR-0040273 at -280 ████████████████████ ██████████████████████████████ *id*. Ex. 44, TRCC-00686613 at -627.

The key number system is a method for indexing the law introduced in 1908. *id.* Ex. 25, TRCC-00608822 at -826. The topics and subtopics in the key number system are dictated by the language of the law. *Id*. Ex. 32, TR-0179830 at -834. Because their purpose for lawyers is to find cases, the terms used are terms lawyers are familiar with. *Id*. Ex. 36 at 203:22-205:21. Topics and subtopics change as the law changes. *Id*. 205:23-206:25. And topics have been added or removed as the law has evolved. *Id*. 202:7-203:10, 204:19-25; Leiter Decl. ¶ 26.

Plaintiffs' copyrighted materials are registered as a compilation. Suppl. Canter Decl. Ex. 39, TR-0034557 at -558. As of June 5, 2019, this compilation included 11,205,374,706 data records. *Id*. Ex. 40, TR-0026188 at-190. As of April 2021, there were ████████ headnotes in Westlaw. *Id*. Ex. 31, Resp. to Interrog. 14 at 7-8. The key number system has 400 topics that cover around 100,000 subtopics. *Id*. Ex. 32, TR-0179830 at -834-36; *id*. Ex. 41, TR-0523445 at -451. As of April 2023, there were over ████████ cases in Westlaw. Resp. to Interrog. 26, at 25.

ARGUMENT

This Court should find that ROSS's use of Westlaw Content is protected by fair use. 17 U.S.C. § 107. ROSS's use of any Westlaw Content was transformative and intermediate. No traces of the Westlaw Content exist on the ROSS platform. As a result, while ROSS may have created a product that competes with Westlaw, it did not create a substitute for Westlaw, because ROSS's technology does not rely on Plaintiffs' expressive content to operate and a consumer cannot access or view Plaintiffs' expressive content within ROSS's platform. ROSS's use of the Westlaw Content was so transformative that not only could one not access any Westlaw Content within ROSS, but one could not reverse-engineer ROSS's outputs to learn what Westlaw Content was used. Marks Decl. ¶ 23. The proper framework for analyzing these facts is thus found in the Supreme Court's recent decision in *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1 (2021), and the Ninth Circuit's decisions in *Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992) and *Sony Computer Ent., Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000).

Plaintiffs argue that *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023) provides the proper framework. But that case has no applicability here. In *Warhol*, the original work and infringing work were the same with differences in limited aspects. *Warhol*, 598 U.S. at 517 (Figure 1, depicting the original image); *id*. at 519 (Figure 3, depicting "Orange Prince," the infringing image); *id*. at 522 (Figure 6, comparing the two works side-by-side). The equivalence to this case would be if ROSS's output was legal opinions with different colored headnotes, synapses, and whatever else falls within the scope of "Westlaw Content."

In contrast, ROSS's use of the Westlaw Content was highly transformative, and ROSS did not use Plaintiffs protected expression. Statement of Facts § IV-VI. Plaintiffs' copyrights are over work that compiles and reports facts. *Id*. § VIII. That copyright is necessarily a thin one. Again, in contrast to *Warhol*. Plaintiffs only real response to this argument relies on the rejected

20

theory of "sweat of the brow," a theory debunked in *Feist*. *Feist*, 499 U.S. at 354-62; 4 NIMMER

ON COPYRIGHT § 3.04[B][1] (2024).

Nor did ROSS take any substantial portion of Plaintiffs' work; the percentage of either

the headnotes or the key number system used was tiny. Finally, Plaintiffs rely on a misplaced

argument for the market. They argue that because ROSS created a competitive product, it must

be a market substitute. But nothing that ROSS did affects the market for Plaintiffs' content. If a

researcher wants that content, he or she must still turn to Westlaw. What ROSS provided was

better search functionality, which copyright law does not protect.

I.   FACTOR ONE FAVORS FAIR USE BECAUSE ROSS'S USE WAS HIGHLY TRANSFORMATIVE

ROSS's use of the Westlaw Content was highly transformative. Plaintiffs argue that

because both ROSS and Plaintiffs use Westlaw Content "to train their respective AI algorithms,"

Brief at 3, ROSS's platform cannot be transformative. Brief at 26-28. That argument misses the

point: ROSS did not use the "expression, meaning or message" of the Westlaw Content. *See*

Google LLC, 593 U.S. at 29 (cleaned up); Statement of Facts § IV-VI. Nor did ROSS even use

the same data as Plaintiffs to create its training data set. *Id*. Plaintiffs also conveniently ignore the

undisputed process that ROSS used to create its training data, which was itself highly

transformative. *Id*. § IV. Lastly, Plaintiffs ignore what ROSS finally used the Westlaw Content

for. *Id*. § V. The Westlaw Content—to the extent used—created a product that removed human

created intermediated content by creating mathematical word relationships so that only the words

of judicial opinions mattered and were returned. *Id*. Plaintiffs used that content to create

algorithms that returned more of the same "Westlaw Content" along with judicial opinions. *Id*.

A.   ROSS Did Not Use Plaintiffs' Expression, Meaning, or Message

ROSS's final product does not contain any Westlaw Content. Statement of Facts § V.

21

When a user enters a query into ROSS, it returns relevant case law. *Id*. It does not return

headnotes, key numbers, or any of the Westlaw Content. *Id*. Nor does it use any Westlaw

Content to search for case law. *Id*. § IV. Instead, it uses a proprietary algorithm based on the

analysis of the abstract linguistic relationships of question-and-answer pairs. *Id*. Any Westlaw

Content used in creating those question-and-answer pairs—and the record shows that not every

pair reflects the use of Westlaw Content—is thoroughly transformed and obliterated in the

creation of the ROSS product. *Id*. § III-IV. That is transformative use. *Id*. §§ IV-VI.

A use is transformative when "the copier's use adds something new, with a further

purpose or different character, altering the copyrighted work with new expression, meaning or

message." *Google LLC*, 593 U.S. at 29 (cleaned up). It is likewise transformative when it is done

to learn or access underlying materials. *See, e.g.*, *Sega,* 977 F.2d at 1522. And a use is

transformative when the original expression is discarded in an intermediate step towards making

a "wholly new product." *Sony*, 203 F.3d at 607; *see also* Pierre N. Leval, *Toward A Fair Use*

Standard, 103 HARV. L. REV. 1105, 1111 (1990). That is what was done here.

ROSS added something new, with a further purpose and different character, by altering

the copyrighted work. It is undisputed that ROSS turned any Westlaw Content into abstract

mathematical relationships about the language used in judicial opinions. Statement of Facts § IV.

To do so, they (through LegalEase,) created the memos. *Id*. § III. ROSS then ███████ certain of

the question-answer pairs that met their quality control guidelines from the memos, which

included multiple classified answers for the same question, into ███████. *Id*. § IV.A. At this

point, any remnants of Plaintiffs' key number system were already eliminated, because certain

question-answer pairs were not extracted into the files. *Id*.

ROSS then randomly selected question-answer pairs from the ██████, using only ███ of

that set for training data, and ▮▮▮ for validation and testing. *Id*. Then, ROSS tokenized and lemmatized the remaining pairs; ROSS removed "stop" words ("to" or "be") (tokenizing) and normalized the grammatical inflection of words in the pairs, (lemmatizing,) so that there would be no distinctions between things like number, tense, or voice. *Id*. Once they had tokenized and lemmatized pairs, ROSS featurized the pairs, meaning that ROSS analyzed the pairs based on twenty-seven statistical, non-expressive, characteristics about the relationships between the words remaining in the pairs. *Id*. § IV.B. Once ROSS derived those features, ROSS taught their technology to match search queries to paragraphs from uncopyrightable case law. *Id*. § IV.C.

The purpose of this effort was like the effort in *Sega* and *Sony* to learn underlying facts that were not copyrightable. *Sega*, 977 F.2d at 1522-23; *Sony*, 203 F.3d at 603-05.

And once done, the "original expression," *i.e.*, any claimed Westlaw Content was discarded. Statement of Facts § IV-V. It never made it onto or into the final actual product—the ROSS search engine. *Id*. Just as Plaintiffs discard their training data so too did ROSS. *Id*. § IV.B.

It is thus undisputed that ROSS transformed any Westlaw Content used in ways the goes far beyond what other courts have held transformative. *See, e.g.*, *Google LLC*, 593 U.S. at 30 (finding fair use even though "Google copied portions of the Sun Java API precisely, and it did so in part for the same reason that Sun created those portions, namely, to enable programmers to call up implementing programs that would accomplish particular tasks."); *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 97 (2d Cir. 2014) (enabling full-text search held transformative); *Authors Guild v. Google, Inc.*, 804 F.3d 202, 207 (2d Cir. 2015) (allowing users to view snippets of books is transformative).

Plaintiffs' cases do not help their argument because they are factually dissimilar. In each of the cases Plaintiffs cite, a consumer of the infringing product could directly access the

23

copyright holder's expressive content. *See, e.g.*, *Worldwide Church of God v. Philadelphia Church of God, Inc.*, 227 F.3d 1110, 1113 (9th Cir. 2000) ("copied [plaintiff's literary work] verbatim"); *Weissmann v. Freeman*, 868 F.2d 1313, 1324 (2d Cir. 1989) (infringing literary work "taken virtually verbatim from prior jointly authored pieces"); *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 914 (2d Cir. 1994) ("photocop[ied] articles in a scientific journal") *Wall Data Inc. v. Los Angeles Cnty. Sheriff's Dep't*, 447 F.3d 769, 774 (9th Cir. 2006) (installed software on more computers than license permitted); *Video Pipeline, Inc. v. Buena Vista Home Ent., Inc.*, 342 F.3d 191, 194 (3d Cir. 2003), *as amended* (Sept. 19, 2003) (video clips taken from a movie); *Oasis Pub. Co. v. W. Pub. Co.*, 924 F. Supp. 918, 926 (D. Minn. 1996) (copied star pagination on CDs); *Hachette Book Grp., Inc. v. Internet Archive*, 115 F.4th 163, 173 (2d Cir. 2024) ("create[d] digital copies of print books and post[ed] those copies on its website where users may access them in full, for free"); *Soc'y of Holy Transfiguration Monastery, Inc. v. Gregory*, 689 F.3d 29, 35 (1st Cir. 2012) (used "verbatim or near-verbatim copies" of plaintiff's texts).

In fact, in four of the cases on which Plaintiffs rely, the infringing party does not argue that its use was transformative. *See Worldwide Church of God*, 227 F.3d at 1177 (arguing only that factor favors fair use because the use is "not commercial or for profit"); *Weissmann*, 868 F.2d at 1324 (arguing only that factor favors fair use due to the nonprofit nature of the use); *Am. Geophysical Union*, 60 F.3d at 920 (defendant acknowledged "the nontransformative nature of its copying"); *Gregory*, 689 F.3d at 60 (defendant "seems to concede that his use of the Works might not have been transformative"). In three other cases, the record "flatly contradict[s]" any evidence of transformation. *W. Pub. Co.*, 924 F. Supp. at 927; *Wall Data Inc.*, 447 F.3d at 778 ("The Sheriff's Department created exact copies of RUMBA's software. It then put those copies to the identical purpose as the original software. Such a use cannot be considered

24

transformative."); *Hachette*, 115 F.4th at 181 (this is a "clear cut" instance of non-transformation because "IA create[d] digital copies of the Works and distribute[d] those copies to its users in full, for free" and IA added nothing new or different to the works). And in *Video Pipeline*, the Court concluded that the video clip previews "lack any significant transformative quality" because they are the same as plaintiff's movie trailers. *Video Pipeline, Inc.*, 342 F.3d at 199-200.

B.   ROSS's Use Was Intermediate

In addition to the extreme transformation done by ROSS, ROSS's use was also transformative because it was intermediate. Intermediate use of copyrighted works to access non-copyrightable information is fair use. *See Sega*, 977 F.2d 1510; *Sony*, 203 F.3d 596; *Atari Games Corp. v. Nintendo of Am. Inc.*, 975 F.2d 832, 843 (Fed. Cir. 1992) ("When the nature of a work requires intermediate copying to understand the ideas and processes in a copyrighted work, that nature supports a fair use for intermediate copying."). The Supreme Court recognizes intermediate copying is fair use as well. *See Google LLC*, 593 U.S. at 22 (citing *Sony* and *Sega*).

To the extent ROSS accessed any of Plaintiffs' copyrights, it only did so to access non-copyrightable information. That non-copyrightable information includes the judicial passages on Westlaw and the non-expressive characteristics of the words in the text. *Id.* § IV.

This is why *Sega* and *Sony* are both directly on point. In *Sega*, the defendant disassembled object code to understand the "ideas and functional concepts embodied in the code." 977 F.2d at 1520. It then used what it learned from those ideas and concepts to create competing games that could be used on the Sega Genesis console. *Id.* at 1522-23. The court found fair use because "[w]here there is good reason for studying or examining the unprotected aspects of a copyrighted computer program, disassembly for purposes of such study or examination constitutes a fair use." *Id.* at 1520. As in *Sega*, ROSS used elements of Plaintiffs' work for a "legitimate, essentially non-exploitative purpose," *id.* at 1523, to study the underlying

words of opinions and non-expressive characteristics about the text—uncopyrightable material.
See id. at 1522; Statement of Facts § IV. That strongly favors a finding of fair use.

In *Sony*, the court went even further. There, it found that repeatedly copying Sony's PlayStation code to understand how it worked was fair use. *Sony*, 203 F.3d at 606-07. It found fair use "notwithstanding the similarity of uses and functions between the Sony PlayStation and the Virtual Game Station," (*id.* at 606) because the Virtual Game Station was a "wholly new product," and did not contain any of the protectible parts of Sony's copyrighted code. *Id.* at 606-07. This "use of the copyrighted material was an intermediate one, and thus was only 'indirect or derivative.'" *Sony*, 203 F.3d at 607 (quoting *Sega*, 977 F.2d at 1522). The court concluded "[w]e are therefore at a loss to see how [plaintiff's] drafting of entirely new object code for its VGS program could not be transformative, despite the similarities in function and screen output." *Sony*, 203 F.3d at 606-07. As in *Sony*, ROSS's product also has, at a general level, similarities in function to Westlaw (both are legal search products). But ROSS's product is "wholly new" compared to Plaintiffs' and just like the Virtual Game System, it does not use any of Plaintiffs' copyrights. Statement of Facts §§ I, IV-VI.

Moreover, as in *Sega* and *Sony*, where the defendants had to access the copyrightable material in order to access the non-copyrightable material, ROSS had to use the text of the question-answer pairs to extract the mathematical features from that text. *See Sony*, 203 F.3d at 603-04; *Sega*, 977 F.2d at 1520-21; Statement of Facts § IV.B.

Plaintiffs try to avoid the intermediate copying cases by arguing that because "ROSS's overall purpose in copying the Westlaw Content in the Bulk Memos was to create an avowed substitute for Westlaw," Brief at 27, ROSS's use could not be transformative. First, the creation of a competitive product does not foreclose fair use generally or a finding that a use is

26

transformative. Acceptance of this argument would utterly foreclose the fair use conclusions in *Sony*, *Sega*, and *Google LLC*.

Second, in the context of factor one, if one uses a work because one wants to use the same expression as the original author, as in *Warhol*, it is much less likely to be transformative. But that is not what happened here. ROSS did not exploit Plaintiffs' intrinsic expressive value; it instead got rid of any expression by Plaintiffs in its entirety to focus on the abstract linguistic relationships of the text to inform its search capabilities. Statement of Facts § IV.

Third, the caselaw does not support Plaintiffs argument. If Plaintiffs were correct that the similarities in overall purpose can defeat fair use, then Google would have lost in *Google LLC*, because Google and Oracle had the same overall purpose for the copyrighted API packages: allowing programmers to call up prewritten software to facilitate application development. *See Google LLC*, 593 U.S. at 30. 2 Live Crew would have lost in *Campbell*, because it had the same overall purpose for the copyrighted lyrics and music as Roy Orbison (creating a song). *Campbell*, 510 U.S. at 572-73. Sony would have lost in the Betamax case because the overall purpose of its tape recorders was the same overall purpose of Universal's broadcasts (having viewers watch the broadcast). *Universal City Studios v. Sony Corp.,* 464 U.S. 417, 419 (1984). Richard Prince and Jeff Koons would have lost in their fair use cases because they had the same overall purposes as Patrick Cariou and Andrea Blanch (creating artistic works). *Cariou v. Prince*, 714 F.3d 694, 698 (2d Cir. 2013); *Blanch v. Koons,* 467 F.3d 244, 246 (2d Cir. 2006). Accolade and Connectix would have lost in their intermediate copying cases because their overall purposes were to build video games or ways to play video games, just like Sega and Sony. *See Sega*, 977 F.2d at 1514; *Sony*, 203 F.3d at 598. And HathiTrust and Google would have lost in their respective cases against Authors Guild because the overall purpose in each case was to make books available to

the public. *HathiTrust*, 755, F.3d at 90-93; *Authors Guild*, 804 F.3d at 206-07.

Plaintiffs also contend that the lesson from the intermediate copying cases is that it is fair use to develop a "compatible product" with another's copyrights. ROSS has not done this. Brief at 30. But compatibility was precisely the thing to which plaintiffs objected and claimed rendered the uses of copyrighted materials in *Sega* and *Sony* objectionable. Indeed, the courts concluded that it was fair use despite compatibility and the fact that the end products were competitive. Therefore, this argument lacks merit. *See Sega*, 977 F.2d at 1522 (the proper focus is Accolade's use of the copyrights to "study the functional requirements" of Sega's product); *Sony*, 203 F.3d at 606 (Connectix had built a product that allows users to play the same game on "a new platform, the personal computer" and that Connectix's product is "wholly new.").

Warhol changes none of this. At issue was a screenprint of the infringed work that look the same as the original except with some additional coloring. *Warhol*, 598 U.S. at 522 (Figure 6, comparing the original and infringing images side-by-side). The fair use argument was that an additional message was added atop the original work but had "no critical bearing on" the original work. *Id*. at 545-47. This was not sufficient, according to the Court. *Id* at 540-41 (rejecting the fair use argument because "'transformative use' would swallow the copyright owner's exclusive right to prepare derivative works" if all a secondary user had to do was add "some new expression, meaning, or message" onto an existing work); *see also* 4 NIMMER ON COPYRIGHT § 13F.05 (2024) ("[u]sing a work for a purpose that exploits its *intrinsic expressive value* thus fails the test [of transformativeness] . . ." (emphasis added). ROSS is not offering something that merely conveys additional meaning or different messaging. And what is offered has no kin to a portrait of a famous singer. Statement of Facts § V-VI. It is obvious that *Sega*, *Sony*, and *Google LLC* read on this case, not a case about an image.

28

In contrast, Westlaw Content is not present in the output of ROSS's AI tool. *Id.* § V. And during ROSS's AI development process, ROSS *removes* any expressive meaning or message that Plaintiffs added to their copyrights. *Id.* § IV. Thus, the Court's concern in *Warhol* that an appropriator will take the original work's meaning or message, concatenate a small addition to the work, and demand transformative use, is absent here because ROSS removed the expression, meaning and message that Plaintiffs gave to their copyrights. *Id.*

By further contrast to *Warhol*, Plaintiffs' claimed uses of their content to train their own search technology did not even rely on the same content as ROSS. Plaintiffs older AI tool, WestSearch, was last trained on only key numbers in 2011. Statement of Facts § VI. And Plaintiffs' newer AI tool, WestSearch Plus, was trained in 2018 on the text of headnotes. *Id.* Plaintiffs now state that this case is about ROSS having infringed the links between headnotes and opinions. *See* Suppl. Canter Decl. Ex. 42, Tr. Pretrial Conference, Aug. 6, 2024 at 34:1-14 ("Our position is they copied the head notes in order to get the pairs with the case text that they go to. . . . This isn't a case about matching language."). This is not the same specific Westlaw Content that Plaintiffs used to train WestSearch or WestSearch Plus.

C.   ROSS's Commercial Use and Alleged Bad Faith Do Not Tip the Scales

The *Google LLC* Court focused on transformative use instead of commercialism because "a finding that copying was not commercial in nature tips the scales in favor of fair use. But the inverse is not true, as many common fair uses are indisputably commercial." *Google LLC*, 593 U.S. at 32. Given the significant transformative use here, the fact that ROSS's use was commercial matters very little. As this Court knows, most of the significant fair use cases would have turned out differently if the fact of a commercial use always weighed greater than transformative use. *See Campbell*, 510 U.S. at 572-73 (commercial use of a song); *Universal City Studios,* 464 U.S. at 419-20 (commercial use of video recorders); *Cariou*, 714 F.3d at 698-

29

03 (commercial use of artwork); *Blanch,* 467 F.3d at 246-47 (same); *Sega*, 977 F.2d at 1514-15

(commercial development of a video game console); *Sony*, 203 F.3d at 598-99 (commercial

development of a video game device); *see also Kienitz v. Sconnie Nation LLC*, 965 F. Supp. 2d

1042, 1049 (W.D. Wis. 2013), *aff'd*, 766 F.3d 756 (7th Cir. 2014) ("[T]he Supreme Court has

made clear that commercial uses are not presumptively unfair.").

Bad faith also has little to no relevance. *See Google LLC*, 593 U.S. at 32. In any event,

ROSS's use of Westlaw Content was not in bad faith. Plaintiffs have no evidence that ROSS

engaged in bad faith. ROSS did not direct LegalEase to use Westlaw and only knew that

LegalEase was using reliable legal research platforms. Statement of Facts § II. LegalEase used

both Westlaw and Lexis to prepare memos. *Id.* And ROSS's attempts in 2015 were years before

the bulk memo project and had no relation to the bulk memo project. *Id.* § IV n.4.

Finally, even were ROSS's prior attempts to access Westlaw relevant, they cannot

foreclose fair use. Google too first sought to license the code at issue in its fair use case. It went

on to use that same code without a license. That was not bad faith sufficient to tip the scales

against Google, even were bad faith an issue. *Google LLC*, 593 U.S. at 8, 32.

II.   FACTOR TWO FAVORS FAIR USE BECAUSE PLAINTIFFS OWN ONLY A THIN COPYRIGHT

Factor two also favors ROSS. The copyrighted work is a factual compilation and is thus

protected by a thin copyright. *See, e.g.*, *S. Credentialing Support Servs., L.L.C. v. Hammond*

Surgical Hosp., L.L.C., 946 F.3d 780, 783 (5th Cir. 2020); *Experian Info. Sols., Inc. v.*

Nationwide Mktg. Servs. Inc., 893 F.3d 1176, 1186 (9th Cir. 2018). Even were the headnotes or

synopses treated as "creative" works, they are still factual, utilitarian, and informational, and thus

removed "from the core of copyright." *See Google LLC*, 593 U.S. at 29.

Plaintiffs argue that the choices made in selecting headnotes and assigning key numbers

are creative. Brief at 33. Assuming this is true, the selections are driven by a utilitarian purpose to allow attorneys to find cases. *See* Statement of Facts § VIII. Whatever creativity exists is constrained by this fact. This is not tabula rasa. A Court has already concluded that the copyright is thin. *Matthew Bender & Co. v. W. Publ'g Co.*, 158 F.3d 693, 707 (2d Cir. 1998).

This makes sense. Westlaw Content's value, in Plaintiffs view, is that it is accurate and comprehensive. Statement of Facts § VIII. Plaintiffs even market that they hire attorneys to draft headnotes to ensure accuracy. *Id.* This Court should thus conclude that "the manifestly factual character of [work] precludes [it] from considering the articles as within the core of the copyright's protective purposes[.]" *Am. Geophysical Union*, 60 F.3d at 925 (cleaned up).

It is undisputed that headnotes are meant to factually report the holdings of the case. Plaintiffs characterize what their attorney-editors do as "identif[ying] and select[ing] the key legal concepts from across the entire paragraph[.]" D.I. 675 at 3; *see also id.* at 5 n.3 (attorney-editors "select[] and arrang[e] the Headnotes, link[] them to selected portion of the opinions, and then classify[] them."). While this may involve some creativity, courts analyzing similar works have consistently found that this factor favors fair use. Newspaper articles, for example, are certainly more creative than headnotes, yet: "[a] number of cases that have analyzed alleged copying of news articles or videotapes of news events have concluded that the second fair use factor weighs in the defendant's favor." *Los Angeles Times v. Free Republic*, 2000 WL 565200, at *15 (C.D. Cal. Apr. 4, 2000) (collecting cases).

Plaintiffs' work is also highly functional. While many cases analyzing factor two often distinguish between factual and creative work, and only discuss functionality in the context of copyrightability, 4 NIMMER ON COPYRIGHT § 13F.06 (2024), courts have found that functionality is also worth considering. *See, e.g.*, *Google LLC*, 593 U.S. at 26-29; *id.* at 50-52 (Thomas, J.,

31

dissenting). When the purpose of a work is to convey factual information, *i.e.*, when its purpose is functional and utilitarian, rather than to "showcase . . . expression," the copyright protection is thin. *Swatch Grp. Mgmt. Servs. Ltd. v. Bloomberg L.P.*, 756 F.3d 73, 89 (2d Cir. 2014).

Plaintiffs' argument also relies on the now-rejected theory of "sweat of the brow." Throughout their briefing, they refer to the "immense creativity *and effort*" that goes into the Westlaw content. Brief at 33 (emphasis added); *see also id.* at 36 (referring to the "investment of significant time and resources" in making Westlaw content; *id.* at 4-5 ("substantial labor and investment."). But the effort expended to make content is irrelevant to copyright law. *Feist*, 499 U.S. at 359-60. The only thing that matters is whether the work is original and creative. *Id.* at 345 ("The *sine qua non* of copyright is originality."); *Utopia Provider Sys., Inc. v. Pro-Med Clinical Sys., L.L.C.*, 596 F.3d 1313, 1319 (11th Cir. 2010) (accord).

Not only is the investment made in a work irrelevant to whether it is copyrightable, it is also irrelevant to factor two. Plaintiffs claim that the Second and Ninth Circuits consider investment under factor two. Brief at 36. But the cases they cite are not good law for this point. Plaintiffs cite *Wall Data Inc. v. Los Angeles Cnty. Sheriff's Dep't*, 447 F.3d 769, 780 (9th Cir. 2006), which relies on *MCA, Inc. v. Wilson*, 677 F.2d 180, 182 (2d Cir. 1981) for this claim. But *MCA* was decided ten years before *Feist*, in which the Supreme Court held that the Second Circuit had "fully repudiated" sweat of the brow as a viable theory of copyright. *Feist*, 499 U.S. at 360; *see also* 4 NIMMER ON COPYRIGHT § 13F.06 (2024) ("in light of the [*Feist decision*] … it is questionable whether [the *MCA*] ruling survives."). ROSS is aware of no post-*Feist* cases other than *Wall Data* holding that a party's investment in its work matters under factor two, nor would it make sense to consider a party's investment. Factor two is concerned with the nature of the copyrighted work, which as argued above, can only be about originality and creativity. *See*

Worth v. Selchow & Righter Co., 827 F.2d 569, 573 (9th Cir. 1987) (collecting cases).

Plaintiffs' copyright covers a factual compilation. Any added creativity is necessarily constrained by its functional purpose, to allow legal researchers to find cases quickly. As a result, Plaintiffs' copyright is thin, and this factor favors ROSS.

III.   FACTOR THREE FAVORS FAIR USE BECAUSE ROSS TOOK A VANISHINGLY SMALL PORTION OF THE WESTLAW CONTENT

Factor three favors ROSS. Under this factor, the inquiry into the "amount and substantiality of the portion used in relation to the copyrighted work as a whole," 17 U.S.C. § 107(3) is both quantitative and qualitative. 4 NIMMER ON COPYRIGHT § 13F.07 (2024). A small portion of a work may still be a substantial portion because "the question of fairness asks what else the parodist did besides go to the heart of the original." *Campbell*, 510 U.S. at 589. And by contrast, "[e]ven entire verbatim reproductions are justifiable where the purpose of the work differs [enough] from the original." *Tresóna Multimedia, Ltd. Liab. Co. v. Burbank High Sch. Vocal Music Ass'n*, 953 F.3d 638, 650 (9th Cir. 2020) (cleaned up). But ROSS took neither a large quantitative portion of the work nor the "heart" of the work, and this factor favors fair use.

ROSS used a tiny percentage of the Westlaw Content. At most, ROSS infringed 21,878 headnotes. *See* Suppl. Canter Decl. Ex. 43, App. B to Suppl. Krien Report Column "Row." This is less than 0.1% of the over ▮▮▮▮▮▮ headnotes. Statement of Facts § VIII. At most, ROSS infringed 5,473 subtopics. *See Id.* Ex. 43, App. B to Suppl. Krein Report Column "WKNS Topic + Key Number"; Declaration of Barbara Fredericksen-Cross, D.I. 688, ¶¶ 23-24. This is 5.473% of the subtopics in the key number system. Statement of Facts § VIII. At most, ROSS copied the 500 cases that it received from LegalEase. Statement of Facts § II. This is less than 0.01% of the ▮▮▮▮▮▮ cases in Westlaw. *Id*. § VIII. And adding up all the headnotes, key number subtopics, and cases, this is just .000444% of the over 11 billion records that Plaintiffs have registered. *Id*.

Plaintiffs nonetheless argue that factor three weighs in their favor because they claim that

ROSS's use of Westlaw Content was "indisputably large." Brief at 38. What this means is never

quantified by Plaintiffs. And as just explained, it is also belied by the numbers.

Also, ROSS extracted factual information from the Westlaw Content, and did not use its

expression. Statement of Facts § IV. This weighs in ROSS's favor on factor three. *Nihon Keizai*

Shimbun, Inc. v. Comline Bus. Data, Inc., 166 F.3d 65, 73 (2d Cir. 1999) ("what is relevant is the

amount and substantiality of the copyrighted *expression* that has been used, not the *factual*

content of the material in the copyrighted works.") (citation omitted) (emphasis in original)).

This minute percentage of a factual, utilitarian, and informational compilation is fair use.

Matthew Bender, 158 F.3d at 678 ("West conceded (immediately prior to trial) [what] would be

permissible under the fair use doctrine, *i.e.*, one to two percent of its published reports of

Supreme Court and court of appeals cases."); *Google LLC*, 593 U.S. at 33-35 (copying 0.4% of

work fair use); *New Era Publ'ns Int'l, ApS v. Carol Publ'g Grp.*, 904 F.2d 152, 158-59 (2d Cir.

1990) (factor favors fair use when defendant used a "small percentage" of copyrighted work.").

Recognizing that ROSS took an infinitesimal portion of their work, Plaintiffs argue that

this factor favors them because ROSS took the "heart" of Westlaw. Brief at 39.[7] Nowhere do

Plaintiffs explain what this means. When case law references "the heart" of a copyrighted work,

it is referring to "what is distinctive and valuable about plaintiff's expression." 4 NIMMER ON

COPYRIGHT § 13F.07 (2024). "When that aspect is in doubt, the most reliable touchstone is, once

again, the risk of substitution." *Id*. Plaintiffs make no attempt to meet this definition.

Plaintiffs' argument is also flawed because it relies on the assumption that its factual

---

[7] Plaintiffs also argue that this factor weighs in Plaintiffs' favor because Plaintiffs contend that factor one weighs in their favor. Brief at 36. ROSS of course disagrees. Argument § I.

compilation has a "heart" the taking of which can be substantial. But "of course, not all works have a heart." Patry on Fair Use § 5:1. Plaintiffs' copyright is a factual compilation; it has little original expression. "Here, by contrast [to *Harper & Row*], there is no objective core of expression in the course materials that can be similarly identified." *NXIVM Corp. v. Ross Inst*., 364 F.3d 471, 481 (2d Cir. 2004) (a work made up of a series of modules had no "core."). As in *NXIVM*, the value of the copyright is in its unitary work, consisting of many parts. *See id.* at 480-81; Statement of Facts § VIII. This kind of copyright can only be substantially taken by a significant quantitative taking, because it has no "heart" within the meaning of this factor.[8]

Additionally, ROSS cannot have taken the "heart" of Plaintiffs' expression. By their own reckoning, the heart of their expression is the annotations on the cases, and the categorization of the cases in a cohesive system. *See* Brief at 39-40. None of that material exists in ROSS. Statement of Facts § IV-VI. One cannot access the key number system by using ROSS. *Id*. § V. One cannot even tell what cases ROSS used to create its training data, or which question-answer pairs were actually derived from. *Id*. In cases where courts have found that a small percentage of a work was nonetheless "substantial," it was because the portion taken would mean that consumers would have no need to buy the original work. *See, e.g.*, *Harper & Row, Publrs. v. Nation Enters*., 471 U.S. 539, 564-65 (1985). In *Harper & Row*, "[t]he portions actually quoted were . . . among the most powerful passages in those chapters." *Id.* at 565. They were chosen because "they qualitatively embodied [President] Ford's distinctive expression." *Id.* What ROSS did is far more analogous to what Google did in *Authors Guild.* There, the defendant made a very small portion of the copyrighted work available to the public, and the court found that this was

---

[8] This principle coheres with the idea that a supersubstantial amount of a compilation must be copied to prove infringement. *See Ross, Brovins & Oehmke, P.C. v. Lexis Nexis Grp.*, 463 F.3d 478, 480 (6th Cir. 2006); *Schoolhouse Inc. v. Anderson,* 275 F.3d 726, 729-30 (8th Cir. 2002).

not a substantial use of plaintiff's copyright. *See Authors Guild*, 804 F.3d at 222; 4 NIMMER ON

COPYRIGHT § 13F.07 n.82 (2024) (discussing *Authors Guild v. Google*).

Because one could not capture any of Plaintiffs' protectable expression by using ROSS,

ROSS's use of the Westlaw Content is closer to *Authors Guild* than to *Harper & Row*. This

factor thus favors ROSS.

IV.   FACTOR FOUR FAVORS FAIR USE BECAUSE ROSS DOES NOT USURP THE MARKET FOR WESTLAW CONTENT

Factor four favors a finding of fair use. Factor four asks whether the alleged infringement

"usurps or substitutes for the market of the original work." *Castle Rock Entm't, Inc. v. Carol*

Publ'g Grp., Inc., 150 F.3d 132, 145 (2d Cir. 1998). Plaintiffs make two arguments as to this

factor. First, they claim that ROSS's search engine usurps the market in legal search engines.

Second, Plaintiffs claim that ROSS usurped the market in AI training data. Neither is valid.

A.   The Legal Search Engine Market is Not Relevant

Plaintiffs first argument depends on a syllogism. They argue that a competitor is one that

produces a product that substitutes for its competition, ROSS was a competitor to Westlaw, and

therefore, ROSS must make a product that substitutes for Westlaw. Brief at 14-15. This type of

reasoning is not recognized by factor four.[9]

Also, Plaintiffs' share of this market is not protected by factor four. "In considering the

fourth factor, our concern is not whether the secondary use suppresses or even destroys the

---

[9] *See* 4 NIMMER ON COPYRIGHT § 13F.08 (2024) ("The inquiry into potential markets gives rise
to a danger of circularity—a potential market, no matter how unlikely, has always been
supplanted in every fair use case, to the extent that defendant, by definition, has made some
actual use of plaintiff's work, which use could, in turn, be defined in terms of the relevant
potential market. In other words, it is always a given that plaintiff suffers a loss of
some potential market if that potential is defined as the theoretical market for licensing the very
use at bar.").

market for the original work or its potential derivatives, but whether the secondary use usurps or substitutes for the market of the original work." *Castle Rock*, 150 F.3d at 145; *Campbell*, 510 U.S. at 591-92. "The only market harms that count are the ones that are caused because the secondary use serves as a substitute for the original, not when the secondary use is transformative …" *Authors Guild*, 755 F.3d at 99; *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630, 643 (4th Cir. 2009) ("An adverse market effect, in and of itself, does not preclude application of the fair use defense."). "The fair use doctrine protects against a republication which offers the copyrighted work in a secondary packaging, where potential customers, having read the secondary work, will no longer be inclined to purchase again something they have already read." *Sundeman v. Seajay Soc'y, Inc.*, 142 F.3d 194, 207 (4th Cir. 1998) (cleaned up).

ROSS does not offer a product that usurps any relevant copyright. ROSS offers an alternative product that a consumer might prefer to Westlaw because ROSS's product provides a different type of research tool than what Westlaw provides. Statement of Facts §§ I, V. "A use that does not materially impair the marketability of the copyrighted work generally will be deemed fair." *Sundeman*, 142 F.3d at 206. ROSS is no more a market substitute for Westlaw than a digitized snipped of text is a substitute for a book, or a parody is a substitute its focus.

ROSS offers nothing that precludes the market that a court would consider under this factor. A legal researcher choosing ROSS might decide not to use Westlaw, but that researcher cannot use ROSS to access, use, or read any copyrighted material on Westlaw. *Id*. § V. "[H]arm caused solely by the fact that the parties compete in the same general market does not by itself equate to harm within the scope of Section 107(4)." Patry on Fair Use § 6:7; *see also, e.g.*, *Lewis Galoob Toys, Inc. v. Nintendo of Am., Inc*., 780 F. Supp. 1283, 1294 (N.D. Cal. 1991) (Plaintiffs

must "demonstrate that use of the Game Genie 'fulfills the demand for the original' of the copyrighted games at issue. *Fisher v. Dees*, 794 F.2d 432, 438 (9th Cir. 1986).").

Plaintiffs' copyright protects tangible expression found within that legal research platform, not the platform itself. *See Arica Inst., Inc. v. Palmer*, 970 F.2d 1067, 1078 (2d Cir. 1992) ("Defendant, after all, is perfectly entitled to create a competing work. The copyright statute simply constrains her from using plaintiff's original expression in doing so."); *Authors Guild*, 804 F.3d at 224 ("the type of loss of sale envisioned above will generally occur in relation to interests that are not protected by the copyright.").

Plaintiffs' claim is really that a copyright protects the form of the work, rather than the work itself. But ROSS is a "substitute" for Westlaw in the same way that a Beatles album is a substitute for a Rolling Stones album. One might (and probably should) prefer the former. Buying the former might even make it impossible for one to purchase the latter, depending on one's financial situation. But that does not mean that the former infringes the latter. It merely means given a choice some consumers will pick it over the latter, which is precisely the kind of choice copyright law is meant to encourage. *See Twentieth Century Music Corp. v. Aiken,* 422 U.S. 151, 156 (1975); *Satava v. Lowry*, 323 F.3d 805, 807 n.1 (9th Cir. 2003).

Put another way, the question is what is it that Plaintiffs sell that is protected by copyright? By their own contention, they offer "highly valuable *content* that helps researchers find and understand the law." Brief at 1 (emphasis added). That content includes headnotes and the key number system. Statement of Facts § VIII. Plaintiffs calculate their copyright damages based on the value of those materials, including lost licensing fees. Suppl. Canter Decl. Ex. 31, Second Supp. Resp. Rog 21. A ROSS customer would not get any of those materials; they simply do not exist in ROSS's platform. Statement of Facts § V. What Plaintiffs attempt to do is

what copyright law expressly forbids: they seek to use their state-issued monopoly on Westlaw

Content to claim a monopoly over the public domain judicial opinions the content accompanies.

Applied Innovations, Inc. v. Regents of Univ. of Minn., 876 F.2d 626, 636 (8th Cir. 1989);

Authors Guild, 804 F.3d at 225 ("The copyright resulting from the Plaintiffs' authorship of their

works does not include an exclusive right to furnish the kind of information about the works that

Google's programs provide to the public."); *cf. Golan v. Gonzales*, 501 F.3d 1179, 1183-84 (10th

Cir. 2007) (discussing copyright protection of works already in the public domain).

> **B.   The AI Training Data Market is Not Relevant, But If It is, ROSS Did Not
>   Implicate that Market**

Plaintiffs' other argument is that they might choose to enter the AI training data market,

and that ROSS's use of Westlaw Content deprives them of licensing revenue. Brief at 20.

Factor four is not an invitation to speculate about future loss. 4 NIMMER ON COPYRIGHT §

13F.08 (2024) ("Insofar as the concern under factor four is market substitution, it would seem

that no harm is cognizable when defendant's utilization neither harms plaintiff's existing market

nor diverts revenues in any potential market it might plausibly enter."). For example, in *Am.*

Geophysical Union, the Second Circuit only agreed that there was a traditional, reasonable, or

likely to be developed market after acknowledging that the 83 plaintiff "publishers" that brought

this case "ha[d] created, primarily through the [Copyright Clearance Center, Inc.], a workable

market for institutional users to obtain licenses for" their copyrights. *See Am. Geophysical*

Union, 60 F.3d at 930; *id.* at 914 (defining "Plaintiffs American Geophysical Union and 82 other

publishers of scientific and technical journals" as the "publishers"). In contrast, Plaintiffs have

done nothing to license their copyrights. They have done the opposite—stating without

equivocation that they will not license their content as AI training data. Statement of Facts § VII.

Moreover, nothing that ROSS has done has stopped or inhibited Plaintiffs from either

entering this market or making money in this market. ROSS did not sell training data and has no desire or intention to do so. *Id*. And when LegalEase used Plaintiffs' platform, in Plaintiffs' view, to create training data, LegalEase paid Plaintiffs. *Id*. § II.

Relatedly, there is no evidence of "substitutionary" harm, which is the only harm that matters for factor four. *See* 4 NIMMER ON COPYRIGHT § 13.08[A] (2024). Anyone who wants to make training data using Westlaw Content still must access Westlaw. Statement of Facts §§ II-III, VIII. Plaintiffs thus have the same incentives under the Copyright Act to produce Westlaw Content as they did before. And for all these reasons, were ROSS's use to become widespread, it would still have no impact on Plaintiffs' market.

Finally, this focus on the AI training data market is legally misplaced. The AI training data was the means to obtain the end result. The end result is the subject of query, not this middle step. Had this intermediate step been the inquiry—had Courts only focused on whether there was a market for the copyrighted materials that were used to create something new—factor four would have gone differently in significant fair use cases. For example, in *Google LLC*, Oracle plainly had a very significant business opportunity to license its source code for use in Android application development—the parties even attempted such negotiations. *See Google LLC*, 593 U.S. at 8. But that fact was irrelevant to the factor four analysis, which weighed in favor of fair use. *See Google LLC*, 593 U.S. at 35-40. Indeed, if Plaintiffs were correct, then the fair use analysis would be decisive on factor four alone and collapse to the question: was the copyrighted materials that the infringing party used valuable, *i.e*., could they be licensed. But this alone is not the question, and courts must balance all four factors.

CONCLUSION

This Court should deny Plaintiffs' motion for summary judgment.

40

Respectfully submitted,

POTTER ANDERSON & CORROON LLP

OF COUNSEL:

Warrington S. Parker III
Joachim B. Steinberg
Jacob Canter
CROWELL & MORING LLP
3 Embarcadero Ctr., 26th Floor
San Francisco, CA 94111
Tel: (415) 986-2800

Keith J. Harrison
Crinesha B. Berry
CROWELL & MORING LLP
1001 Pennsylvania Avenue NW
Washington, DC 20004
Tel: (202) 624-2500

By: /s/ David E. Moore
      David E. Moore (#3983)
      Bindu A. Palapura (#5370)
      Andrew L. Brown (#6766)
      Hercules Plaza, 6th Floor
      1313 N. Market Street
      Wilmington, DE 19801
      Tel: (302) 984-6000
      dmoore@potteranderson.com
      bpalapura@potteranderson.com
      abrown@potteranderson.com

*Attorneys for Defendant/Counterclaimant
ROSS Intelligence, Inc.*

Dated: October 30, 2024
11852645 / 20516.00001

 Public Version Dated: November 6, 2024

41